Preliminary Findings from Year 1 Pilot Implementation of the Arizona Department of Education (ADE) Model Teacher Evaluation Process

Arizona Department of Education

West Comprehensive Center

Regional Educational Laboratory West

July 22, 2013

Learning Together

- Many individuals and organizations partnering to learn together about pilot implementation

  - Arizona Department of Education (ADE)

  - 10 Pilot/Partner LEAs

  - West Comprehensive Center (WCC)

  - Regional Educational Laboratory West (REL West)

- Preliminary findings developed collaboratively

- More extensive reporting on pilot year 1 to come from ADE and REL West

Year 1 Data Collection: Pilot LEAs

- Two rounds of focus groups/interviews (February & May)

  - Round 1: 46 teachers, 13 principals

  - Round 2: 35 teachers (5 new), 11 principals (0 new)

- Online surveys emailed to participating teachers in May

  - 165 participating pilot teachers responded

- Mix of grade spans (elementary/secondary) & experience levels represented in focus group/interview, survey data

Overall Findings

Preliminary Findings on Overall ADE Model

- In interviews, principals called the new system less biased than previous systems & generally believed that it resulted in higher quality feedback for teachers.

- Despite many focus group teachers reporting frustrations with ratings lower than expected:
    - 52% of surveyed teachers agreed that the new teacher evaluation process is fair.
    - 45% of surveyed teachers felt new process is an improvement over their prior teacher evaluations.

Preliminary Findings on Training/Communication

- 63% of surveyed teachers agreed that "the criteria on which I was evaluated were made clear to me".

- 55% agreed that their training was adequate for them to effectively participate.

  - Some focus group teachers reported difficulty using *Teachscape* online & hoped for more training on the use of the website.

Preliminary Findings on Training/Communication

- Principals felt well trained to assess Danielson Framework Domains 2 & 3, less prepared to assess Domains 1 & 4.

- Some communication disconnects within LEAs & schools

- Key remaining question areas for participants:

  - Uncertainty around component scoring

  - Equity issues across student populations

  - Among principals: Coaching conversations with teachers

Preliminary Finding: Time Issues

- All participants cited labor intensiveness, time burdens involved with the new, more thorough observations (e.g., new forms, multiple domains, more evidence).

- Time estimates from principals in one focus group:

    - Approximately 3 hours per observation cycle

    - Round of observations for 20 teachers takes about one month

- *Teachscape* technology helped efficiency, but process is still lengthy.

Preliminary Finding: Technology Issues

- Participants had some difficulties with *Teachscape* system, particularly early in implementation.

  - Logging on/navigating the system

  - Automated reminders sent to spam folders

  - Problems finding and uploading documents

- Many of these issues were resolved as the pilot progressed.

Preliminary Results from End-of-Year Surveys of Teachers:

The new ADE model teacher evaluation process has...

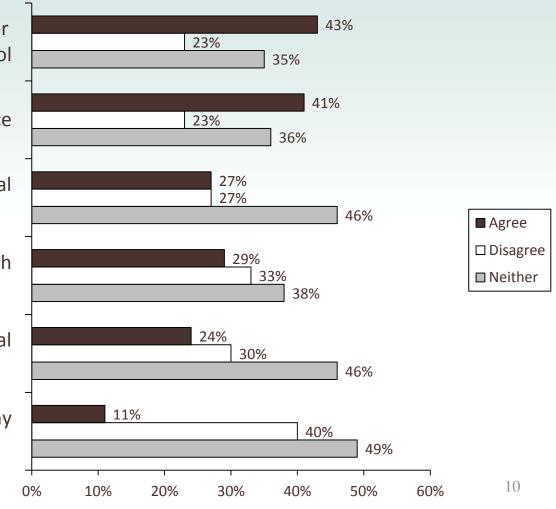Provided a common language for professional practice in my school
- Agree: 43%
- Disagree: 23%
- Neither: 35%

Led me to improve my instructional practice
- Agree: 41%
- Disagree: 23%
- Neither: 36%

Improved the quality of my instructional interactions with my administrator
- Agree: 27%
- Disagree: 27%
- Neither: 46%

Helped me engage in professional growth opportunities targeted to my needs
- Agree: 29%
- Disagree: 33%
- Neither: 38%

Improved the quality of my instructional interactions with my colleagues
- Agree: 24%
- Disagree: 30%
- Neither: 46%

Improved the climate and culture in my school
- Agree: 11%
- Disagree: 40%
- Neither: 49%

Legend:
- ■ Agree
- □ Disagree
- ■ Neither

Other Preliminary Findings from Focus Groups

- To many principals, Danielson Framework provided clear definition of effective teaching & what to look for:

    - Principals thought conversations with teachers were more focused & in-depth (thanks to the Framework) & collaborative, particularly after pre-conferences.

    - But most focus group teachers reported that the new process had <u>not</u> changed conversations with their administrators

- Some participants felt overwhelmed, noting stress & agitation among teachers at the site, particularly after summative ratings were shared.

Planned ADE Response: Training/Communication, Time Management, Technology

- Time Management Support –
    Wednesday Session from 10:45 – 12:15

- HB2500 – Waiver of 2nd Observation (with local board approval)

- Principal Resource Guide

- Teachscape REFLECT & Webinars

- Calendar of Events

- ADE-to-Pilot LEA Communication

Teaching Performance: Observations & Conferencing

(60 points, 50% of total evaluation)

Logistics of Observations/Conferences

- 83% of responding teachers reported 2-3 formal classroom observations
  - But informal observations varied:
    - One informal observation: 30%
    - 2-3 informal observations: 36%
    - 4+ informal observations: 34%

- 71% of respondents reported 2-3 pre-conferences, 63% reported having 2-3 post-conferences
  - These conferences generally spanned less than an hour:
    - < 15 mins (20%)
    - 15-30 mins (45%)
    - 31-45 mins (25%)

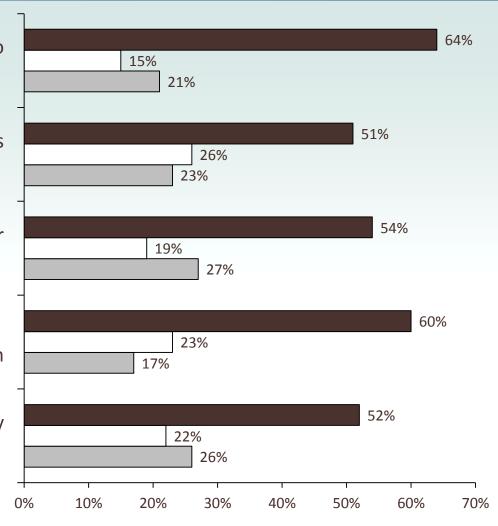# Preliminary Results from End-of-Year Surveys of Teachers: Observations & Conferences

I have confidence in my evaluator's ability to accurately rate my instructional practice
- Agree: 64%
- Disagree: 15%
- Neither: 21%

The number of formal classroom observations was adequate to assess my performance
- Agree: 51%
- Disagree: 26%
- Neither: 23%

Pre-observation conference(s) fully prepared me for what to expect during my observation(s)
- Agree: 54%
- Disagree: 19%
- Neither: 27%

Post-observation conference(s) provided meaningful feedback on how to improve instruction
- Agree: 60%
- Disagree: 23%
- Neither: 17%

My post-observation conference(s) helped identify needs for my professional growth
- Agree: 52%
- Disagree: 22%
- Neither: 26%

Legend: ■ Agree □ Disagree ■ Neither

Focus Groups: Observations/Conferences

- Perceptions of Danielson Framework-based observations were more varied in focus groups/interviews:

  - <u>Positive aspects</u>: Accurate, consistent, reasonable, helpful, specific, evidence-based, objective (less biased)

  - <u>Negative aspects</u>: Time consuming, inflexible, lack of relevant content expertise among observers, too easy to prepare for/script/manipulate (need for additional informal observations)

Preliminary Findings: Summative Conferences & Ratings

- 55% of responding teachers agreed that summative performance classification accurately reflected their overall performance

- In focus groups, the fairness & accuracy of final ratings was a common topic
  - Some principals expressed concerns that certain components unfairly pulled teachers' ratings down
  - Teachers expressed concerns about the difficulty of achieving a Distinguished rating on the Danielson Framework

Preliminary Findings:
Summative Conferences & Ratings

- Some teachers found the post-conference feedback to be thorough, reflective, collaborative, evidence-based & helpful for improving practice.

- Others found it to be a negative experience, seeing it as unfair, not useful, or even insulting to them.

Planned ADE Response: Observations & Conferences

- Leading Instructional Conversations Support –
Wednesday Session from 9:00 – 10:30

- Integration with Arizona's Common Core Standards –
Tuesday Session from 9:00 – 12:00

- Principal Resource Guide

- Teachscape FOCUS (Proficiency System)

- Teachscape LEARN (New)

- Calendar of Events

Student Surveys, Parent Surveys, Peer Review

(20 points, 17% of total evaluation)

Teacher Survey Results: Stakeholder Surveys

- Surveyed teachers were pessimistic about student & parent surveys providing an accurate assessment of their teaching performance.

  - Proportions of teachers responding that the following can assess their performance with moderate/high accuracy:

    - Student surveys: 50%

    - Parent surveys: 46%

- More optimism expressed about the potential for peer review:

  - 69% of surveyed teachers indicated that peer teacher surveys can provide an accurate assessment of their performance.

Focus Group Results: Stakeholder Surveys

- Peer review, student surveys & parent surveys were a major concern in focus groups/interviews at all pilot LEAs.
  - Common student/parent survey concerns:
    - Logistical difficulties with administration (computer/Internet access)
    - Reliability/validity of results: low response rates, overly subjective, problems "assigning" students/teachers, age/maturity of students
  - Common peer review concerns:
    - Confidentiality (some printed forms)
    - Some questions were difficult to answer knowledgeably (e.g., professional organizations), little useful feedback
    - Some reviewers assigned by principals, others picked by teachers
    - Collegiality issues/tensions/discomfort with process

Planned ADE Response: Stakeholder Surveys

- Enhanced Instructions and Administration Support

- Modified Deployment Requirements

- Updated Point Allocation Tables

- Modifications to Peer Review Questions

- Principal Resource Guide

Student Academic Progress

(40 points, 33% of total evaluation)

Teacher Survey Results: Student Test Data

- Surveyed teachers were generally optimistic that student test data can provide an accurate assessment of their teaching performance.

  - Proportions of teachers responding that the following can assess their performance with moderate/high accuracy:

    - SLOs established through consultation with principal: 64%

    - Standardized school-wide test scores: 60%

    - Standardized test scores from their classroom(s) of students this year: 57%

Focus Group Results: Student Test Data

- Focus group teachers expressed confusion/concerns about student achievement component, citing:

  - Fairness issues between Group A/B teachers (different criteria) and/or those with differing student populations

  - Use of lagged test data

  - Potential for new teachers to receive higher ratings (no prior data)

- Some principals unsure about how to interpret/use data tables (though discussions with ADE staff helped)

- Additional guidance on SLOs sought

Planned ADE Response: Student Test Data

- Updated Point Allocation Tables

- Student Learning Objectives

- Principal Resource Guide